Virology **300,** 316–331 (2002)
doi:10.1006/viro.2002.1535

Evidence for a Conserved Polydnavirus Gene Family: Ichnovirus Homologs of the CsIV *Repeat Element* Genes

A.-N. Volkoff,*[1,2] C. Béliveau,‡[1] J. Rocher,* R. Hilgarth,† A. Levasseur,‡ M. Duonor-Cérutti,* M. Cusson,‡ and B. A. Webb†

*I.N.R.A., Laboratoire de Pathologie Comparée, UMR 5087 INRA/CNRS/Université de Montpellier II, 30380 St Christol-les-Alès, France; †Department of Entomology, University of Kentucky, Lexington, Kentucky 40546-0091; ‡Laurentian Forestry Centre, Canadian Forest Service, Natural Resource Canada, P.O. Box 3800, Sainte-Foy, Quebec, G1V 4C7 Canada

Received January 17, 2002; returned to author for revision April 1, 2002; accepted April 3, 2002

In *Campoletis sonorensis* Ichnovirus (CsIV), the *repeat element* genes constitute a gene family of 28 members. In the present work, we document the presence of members of this gene family in two additional ichnoviruses, *Hyposoter didymator* Ichnovirus (HdIV) and *Tranosema rostrale* Ichnovirus (TrIV). Two *repeat element* genes, representing at least one functional gene, were identified in TrIV, whereas HdIV was found to contain at least three such genes. In both HdIV and TrIV, the known *repeat element* genes are encoded on single genome segments, with hybridization studies suggesting the presence of other, related but as yet uncharacterized genes. The HdIV and TrIV *repeat element* genes are all transcribed in infected caterpillars, although differences exist among genes in levels and in tissue specificity of expression. A heuristic tree was generated indicating that the *repeat element* genes are more similar within a species of wasp than between species, with TrIV genes being more closely related to the CsIV than to the HdIV genes. These results suggest that the most significant duplication, divergence, and expansion of the *repeat element* genes occurred after speciation. The finding that *repeat element* genes form an interspecific family within the genus *Ichnovirus* supports the view that the proteins they encode play an important role in ichnovirus biology.

Key Words: ichnovirus; *repeat element* gene family; transcription in lepidopteran host; phylogenetic analysis.

INTRODUCTION

Polydnaviruses are unusual dsDNA viruses associated with parasitic wasps belonging to the families *Ichneumonidae* and *Braconidae.* The segmented polydnavirus genome is maintained as a provirus integrated in wasp chromosomes. Viral replication is restricted to the wasp ovary with viral particles accumulating in the lateral oviducts until injection into lepidopteran hosts during oviposition. Viral replication is not detected in infected lepidopteran cells but viral genes are expressed. Polydnavirus gene products are responsible for several of the physiological modifications observed in parasitized caterpillars, such as immune suppression (reviewed in Beckage, 1998), alteration in the levels of developmental hormones (reviewed in Cusson *et al.,* 2000), and inhibition of host protein translation in the fat body (Shelby *et al.,* 1998; Shelby and Webb, 1997). For several polydnavirus species, genes expressed in the parasitized lepidopteran larva (i.e., genes that are putatively involved in altering host physiology) have now been identified (Béliveau *et al.,* 2000; Johner *et al.,* 1999; Trudeau *et al.,* 2000; Varricchio *et al.,* 1999; Webb, 1998; Yamanaka *et al.,* 1996). Interestingly, within a given polydnavirus, some expressed genes display extensive sequence similarities, suggesting that they are members of viral gene families. For example, a glycine- and proline-rich gene family has been identified in *Hyposoter didymator* Ichnovirus (HdIV) (Volkoff *et al.,* 1999), whereas the genome of *Campoletis sonorensis* Ichnovirus (CsIV) has been shown to harbor at least two gene families, the *cys-motif* genes (Cui *et al.,* 2000; Dib-Hajj *et al.,* 1993) and the *repeat element* genes (Hilgarth and Webb, 2002a,b). Thus, it appears that gene families are a common, perhaps predominant, feature of polydnavirus genomes.

The *repeat element* genes, recognized by the presence of an imperfectly conserved ~540-bp repeated element, were first described by Theilmann and Summers (1987) in CsIV. Hilgarth and Webb (2002b) described a total of 28 known and predicted *repeat element* genes from the CsIV genome, with repeat sequences arranged singly or in direct tandem arrays. The predicted CsIV *repeat element* genes are distributed among 10 different genome segments with discernable evolutionary distinctions between single-repeat and multiple-repeat containing genes. Five *repeat element* genes are known to be expressed in the lepidopteran host (Theilmann and Summers, 1988), whereas at least three are

[1] These authors contributed equally to this work.

[2] To whom correspondence and reprint requests should be addressed. E-mail: volkoff@ensam.inra.fr.

0042-6822/02 $35.00

also transcribed in the wasp host, but evidently at lower levels (Hilgarth and Webb, 2002a). These *repeat element* genes encode proteins with highly significant sequence similarity. Although the function of these genes has not yet been elucidated, their high abundance in the CsIV genome suggests that they may play an important role.

Here, we report on the identification of *repeat element* genes in two additional ichnoviruses, HdIV and the *Tranosema rostrale* Ichnovirus (TrIV), a finding supporting the view that the *repeat element* proteins play a critical role in ichnovirus biology. Two *repeat elements,* representing at least one functional gene, have been identified in TrIV, while HdIV contains at least three *repeat element* genes. In TrIV, the described *repeat element* genes are encoded on a single genome segment while in HdIV, these related genes are found on closely related genome segments. The findings reported herein reveal the existence of a conserved polydnavirus gene/protein family present in multiple species. This gene family provides clear genetic evidence of a relationship among ichnovirus species and may help to elucidate the phylogenetic relationships between ichnoviruses and other viruses/organisms.

RESULTS

HdIV genome segment E encodes three putative *repeat element* genes

Polydnaviruses can infect insect cells in culture, leading to the production of viral transcripts in a manner similar to that observed in parasitized lepidopteran larvae. Using cDNAs generated from such transcripts, we were able to identify *repeat element* genes in HdIV. Thus, among several polydnaviral cDNAs corresponding to mRNAs extracted from HdIV-infected Sf9 cells, 24 h after infection (Volkoff *et al.,* 1999), one (C28) generated four hybridization signals when used as a probe against undigested HdIV DNA in Southern blot analysis. Two of these bands were identified as the superhelical and open-circular forms of a 4.6 kb HdIV genome segment designated E (Fig. 1A, lanes nd; for the picture shown, the probe was segment E but identical results were obtained with C28). Segment E was PCR-amplified using C28-specific primers, HdCS3 and HdCS5, and then cloned and sequenced. The HdIV genome segment E is 4644 bp long and contains three putative open reading frames (ORF), including the one corresponding to the C28 cDNA (GenBank Accession No. AF364055; Fig. 1C).

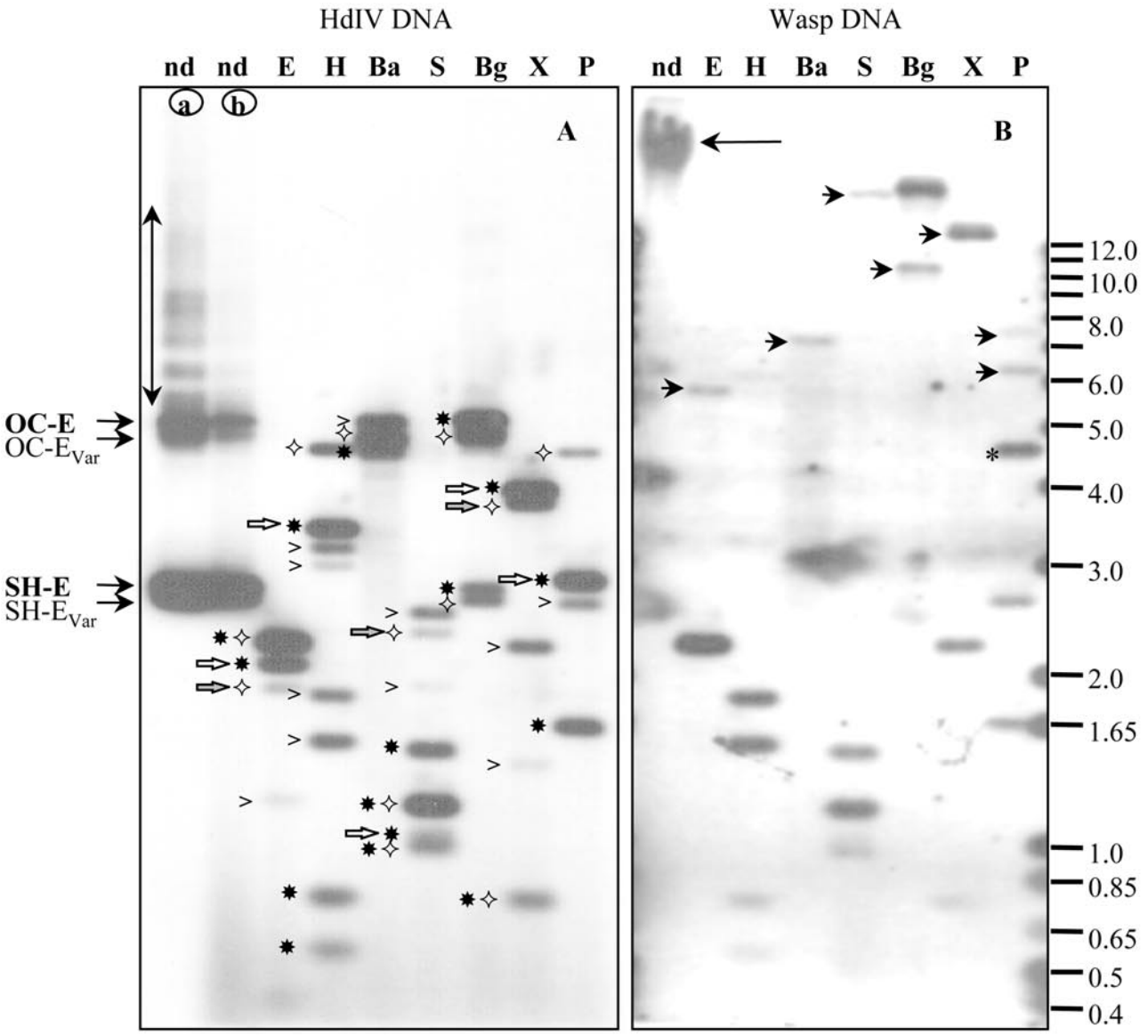
Dot-plot analysis (Lasergene; data not shown) of the segment E sequence revealed that the predicted ORFs were related. In a BlastX search through nonredundant databases, the predicted translation products of all three ORFs showed significant similarity to the protein encoded by the CsIV BHv0.9 *repeat element* gene (GenBank Accession No. M23437; named CsB*rep1* in the present article). In view of this similarity, the three putative genes were designated HdE*rep1* (corresponding to C28), HdE*rep2,* and HdE*rep3* (Fig. 1C), with respective BlastX values of $4e^{-24}$, $4e^{-19}$, and $5e^{-26}$ for similarity with CsIV CsB*rep1.*

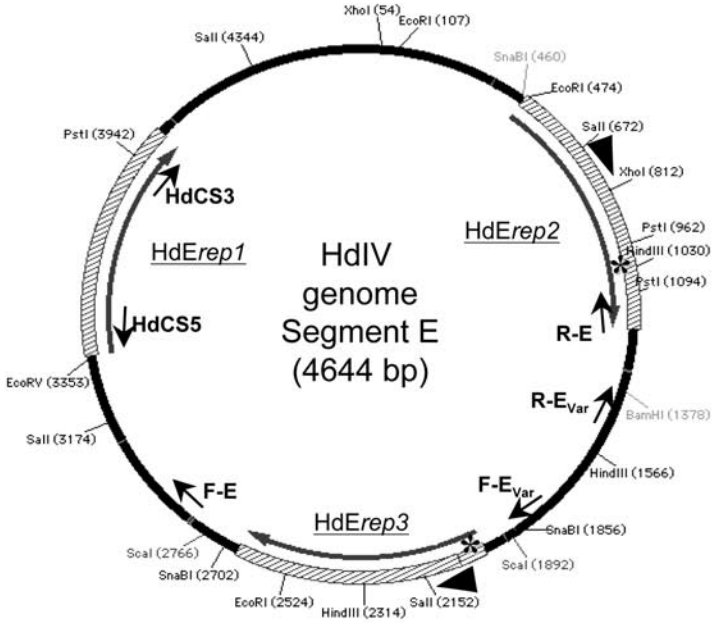
Only 45% of the genome segment E corresponds to coding sequence. The HdE*rep1,* HdE*rep2,* and HdE*rep3* ORFs are 678, 735, and 693 nt long, respectively; they encode putative proteins of 225, 244, and 230 amino acids, with predicted molecular weights of 26, 28, and 27 kDa, respectively. The ATG codons are in a context favorable to translation initiation, according to Kozak (1989). No consensus for TATA box or CAAT box is found upstream of the three ORFs. In the 3′UTR, a polyadenylation signal AATAAA is present for all three ORFs. Intron sequences are not predicted, which was confirmed by sequencing of the 5′ and 3′RACE products (see below). The three ORFs display 32% overall nucleotide identity. Compared in pairs, overall identities are 49, 48, and 45% for HdE*rep1/rep2,* HdE*rep2/rep3,* and HdE*rep1/rep3,* respectively. By analogy to the CsB*rep1* gene, the ~540-bp repeat starts at position +174 (relative to the start codon) for the HdE*rep1* and HdE*rep2* genes, at position +151 for the HdE*rep3* gene. In the 5′ region upstream the single repeat, the three ORFs share 22% identical nucleotides, while in the repeat, 35% of the nucleotides are identical, with 42% identical nucleotides over the first 450 nt. The highest pairwise identities, from 54 to 57%, are also seen in the first 450 nt of the ~540-bp repeat.

A variant of HdIV genome segment E also encodes *repeat element* genes

The other two hybridization signals to the HdE*rep1* (C28) probe (Fig. 1A) corresponded to a smaller viral segment, which was also PCR-amplified using the primers HdCS3 and HdCS5. Sequence analysis of the amplimer indicated that it was identical to segment E except for a 876-bp region, termed "variable region" (flanked by asterisks on Fig. 1C). This variable region was flanked by a 103-nt imperfectly conserved direct repeat (78% identity) present in both segments (arrowheads on diagram, alignment on the right, Fig. 1C). This second segment was designated "E variant" or E_{Var} (GenBank Accession No. AF364056). Sequence comparison of the segment E_{Var} variable region with the corresponding region of segment E (1049 nt) showed only 43.7% nucleotide identity due to several deletions in E_{Var} (difference of 177 nt between the two segments) and to several nucleotide substitutions, suggestive of point mutations. The sequence divergence resulted in intersegmental differences in the ORFs of HdE*rep2* and HdE*rep3.* The 3′ region of HdE_{Var}*rep2* was shorter due to a displacement of the stop codon, while the 5′ region of HdE_{Var}*rep3* differed by several amino acids.

The restriction fragments expected from the digestion of segments E and E_{Var} were detected by Southern analy-





C

```
662    tgtggcgaat gtcgactcat cggttcacag
       ||||| | || |||||||| | | ||||| |
2042   tgtggaggat gtcgactcgt aagatcacaa

692    ctacatttct caacgggaaa ccattggtga
       | | ||| | |||||| || ||| | | ||
2072   caaagtttat caacggagaa ccaatagaga

722    tccgatacaa ctacgacccc tcaagattag
       ||   ||||| |||||  ||   |||| |||
2102   tcgagtacaa ttacgatccg ggaagaatag

752    aagaagaacg tgt 764
       ||||||||||| |||
2132   aagaagaacg tgt 2144
```

sis of digested HdIV DNA, using segment E as a probe (Fig. 1A; ✱, E; ✧, E_{Var}). However, additional hybridization signals were observed, indicating that identical and/or similar sequences exist in other segments (Fig. 1A, arrowheads). Prolonged exposure of the film to the membrane revealed the presence of weaker hybridization signals involving genome segments of molecular mass greater than that of segment E (Fig. 1A; left-most lane, double-headed arrow). These data suggest that additional, as yet uncharacterized, *repeat element* genes exist in the HdIV genome.

Polymorphism and integration of HdIV genome segment E

Sites at which polydnavirus DNA segments are integrated into wasp DNA are marked by direct terminal repeats of variable length and homology (Fleming and Summers, 1991; Gruber *et al.,* 1996; Savary *et al.,* 1997). Variants of genome segments can be produced as a result of alternate excision of variants of the terminal repeat (Fleming and Summers, 1991). Alternatively, segment variants could exist at different loci in the genome and represent segments that have been duplicated recently. The latter situation could lead to duplication of *repeat element* genes and allow for genetic and functional divergence.

In an effort to determine which of these two possibilities accounted for the observed polymorphism of segment E, hybridization and PCR analyses were performed. Both methods showed that the point of segment E chromosomal insertion lies outside the variable region, ruling out the hypothesis that this region would correspond to the chromosomal integration site for both genome segments. In Southern blot analyses, some restriction fragments obtained with circular viral DNA were replaced by hybridization signals of lower mobility following digestion of wasp genomic DNA (Fig. 1B, filled arrowheads), as would be expected of an integrated, linearized segment. Analysis of the missing fragments (white arrow heads in Fig. 1A) pointed to the region between the *Eco*RI and *Hin*dIII restriction sites, within ORF HdE*rep3,* as the point of segment E chromosomal insertion. This was confirmed by the analysis of the PCR products generated from genomic wasp DNA, using two different sets of primers: whereas the use of primers F_1/R_1, designed to amplify the variable region immediately upstream of the *Hin*dIII site (see diagram, Fig. 2C) resulted in an amplimer of the expected size, the use of primers F_1/R_2, designed to amplify the putative integration site from a point downstream of the *Eco*RI site, resulted in no amplification product (Fig. 2A).

With respect to segment E_{Var}, Southern blot analysis of wasp genomic DNA pointed to the region between the *Eco*RI and *Xho*I restriction sites as the region of segment E_{Var} chromosomal insertion. Although this region encompassed the *Eco*RI-*Hin*dIII restriction region of segment E chromosomal insertion, the absence within the variable region of restriction sites for the enzymes employed here precludes the drawing of definitive conclusions about a same excision site for both segments. Interpretation was complicated by the unexpected presence of a *Pst*I restriction fragment (Fig. 1B, ∗) on the wasp genomic DNA blot, with a molecular mass apparently identical to that of the entire segment E_{Var}, as seen on the viral DNA blot (Fig. 1A). This band likely reflects the presence of another *Pst*I restriction site, within the wasp chromosome, located about 4.5 kb from the E_{Var} *Pst*I restriction site. Available PCR evidence indicates that the point of segment E_{Var} chromosomal insertion differs from that of segment E. Indeed, no segment E_{Var}-related fragments were PCR amplified from genomic wasp with either set of primers (F_1/R_1, common to E and E_{Var}, or F1/R-E_{Var}, specific to E_{Var}; Fig. 2B), indicating that the point of E_{Var} chromosomal insertion possibly lies between the *Xho*I restriction site and the R-E_{Var} primer. These results suggest that the two segments are integrated at different loci; however, a study of the flanking sequences will be necessary to confirm this conclusion.

TrIV genome segment F encodes two putative *repeat element* genes

Identification of TrIV *repeat element* genes was accomplished using cDNAs generated from mRNAs obtained from infected host caterpillars, as opposed to cells in culture. In a previous study (Béliveau *et al.,* 2000), two TrIV

FIG. 1. HdIV genome segment E contains three *repeat element* genes and is integrated in wasp DNA. Southern blot analysis of HdIV (A) and wasp genomic DNA (B), either nondigested (nd) or following digestion with *Eco*RI (E), *Hin*dIII (H), *Bam*HI (Ba), *Sal*I (S), *Bgl*II (Bg), *Xho*I (X), and *Pst*I (P), using segment E as a probe. DNA quantities loaded per lane were 4 and 2 μg, for wasps and HdIV, respectively (except for HdIV lane "nd-(a)," where 4 μg were loaded). Exposure periods: 24 h for HdIV DNA, 48 h for wasp DNA. OC-E and OC-E_{Var}: segments E and E_{Var} in their open-circular topology; SH-E and SH-E_{Var}: segments E and E_{Var} in their super-helical topology. Symbols in panel A: ✱, segment E-related fragments; ✧, segment E_{Var}-related fragments; >, supernumerary fragments; double-headed vertical arrow, high molecular weight HdIV segments that generated a hybridization signal when more DNA was loaded; white arrows, fragments detected only in digests of HdIV circular DNA. Symbols in B: arrow, undigested male chromosomal DNA; black arrowheads, high molecular weight restriction fragments detected only in digests of male wasp DNA; ∗, restriction fragment of the same apparent size as linearized segment E_{Var} seen in A (see text for details). Numbers on the right-hand side of B: linear DNA molecular weight markers (kbp). (C) Restriction map of HdIV segment E showing the position of the three *repeat element* ORFs. Boxes represent potential ORFs, beginning with an ATG start codon, arrows indicating the 5′ → 3′ sense. The "variable" region, for which sequences differ between segments E and E_{Var}, is flanked by asterisks. On the right, nucleotide alignment of the repeated sequence (corresponding to regions marked by arrowheads in diagram) flanking the variable region. PCR primer positions are indicated with arrows.

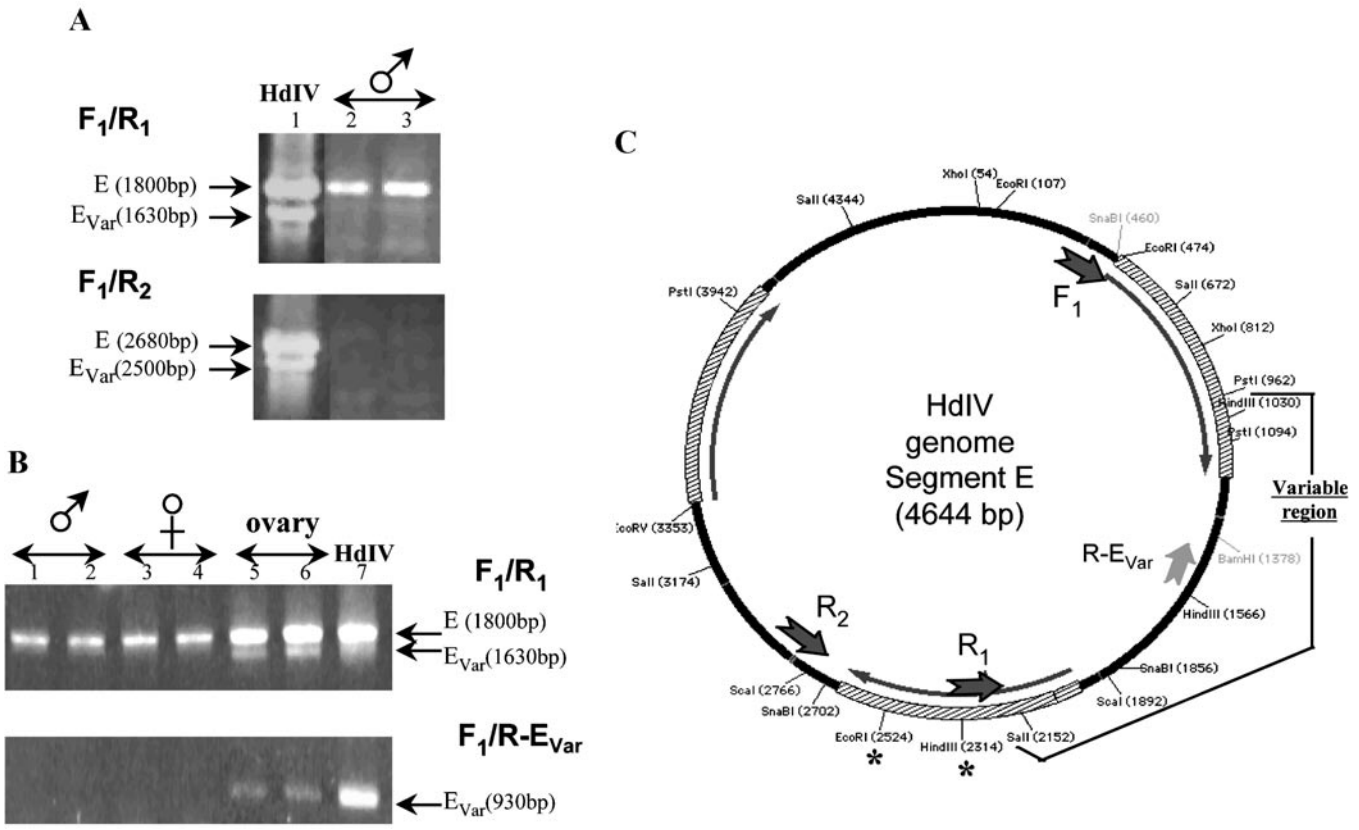


FIG. 2. Integration of HdIV genome segments E and E_{Var}. (A) PCR products obtained from HdIV and wasp DNA using two sets of primers common to both segments E and E_{Var}:F_1/R_1 and F_1/R_2 (represented in diagram C). Sizes of the PCR products are indicated on the left. (B) PCR products obtained from wasp DNA (individual males, females, and ovaries) using one set of primers common to both segments E and E_{Var} (F_1/R_1) and one set of primers specific to segment E_{Var} ($F_1/R\text{-}E_{Var}$) (see diagram C) compared to products obtained from viral DNA. Products were analyzed on ethidium bromide stained gels. Sizes of the PCR products are indicated on the right. (C) Sequences used to design PCR primers for segment E are indicated with arrows. Asterisks indicate the integration region predicted for segment E, between *Eco*RI and *Hin*dIII restriction sites.

genomic libraries were screened with ^{32}P-labeled DNAs complementary to RNAs isolated from TrIV-infected *Choristoneura fumiferana* larvae. Only one of the four families of positive clones obtained was initially sequenced; recently, we analyzed the remaining clones, and one had a restriction fragment that hybridized to the probe initially used for library screening (data not shown). This fragment was sequenced and found to contain a gene encoding a 230 amino acid protein displaying significant similarity to CsB*rep1,* with a BLASTP value of $9e^{-16}$. This gene was designated TrF*rep1* (=*TrV3* in Cusson *et al.,* 2001).

Southern blot analysis (Fig. 3A) of undigested TrIV DNA, using a TrF*rep1*-specific probe, revealed two strong hybridization signals identified as the superhelical and open-circular forms of genome segment F (compare with EtBr-stained gel; Cusson *et al.,* 1998). Additional, weaker signals were detected when the hybridization was conducted at lower stringency (Fig. 3A), suggesting that other genome segments could contain TrF*rep1*-related sequences.

An estimation of the size of the *Sph*I clone carrying the TrF*rep1* gene (~8000 bp) combined with the pattern of bands obtained by Southern analysis (see below) suggested that it may represent the entire segment F. Overlapping restriction fragments representing the entire clone were thus subjected to sequence analysis (GenBank Accession No. AF421353). The latter revealed the existence of one additional *repeat element* related DNA region on segment F, provisionally designated TrF*rep2* (BLASTP value of $7e^{-68}$ for similarity to TrF*rep1;* Figs. 3B and 6) encompassing a 521-nt fragment (nt 3630–4151) displaying 77% identity to the corresponding fragment (nt 4911–5432) in TrF*rep1.* However, its organization called into question its status as a functional gene: the portion over which significant similarity with TrF*rep1* was observed extended beyond (upstream) the putative ATG start codon. If this is a functional gene, one possibility is that it generates a protein that is truncated, relative to TrF*rep1.* Alternatively, it could be a spliced gene, although the described *repeat element* genes of CsIV (Theilmann and Summers, 1988; Hilgarth and Webb, 2002a) and of HdIV are not spliced. In addition to the two *repeat element* related sequences, six small putative ORFs were identified on segment F, none of which showed significant similarity to *repeat element* or other known PDV genes; whether they are part of functional genes remains to be determined.

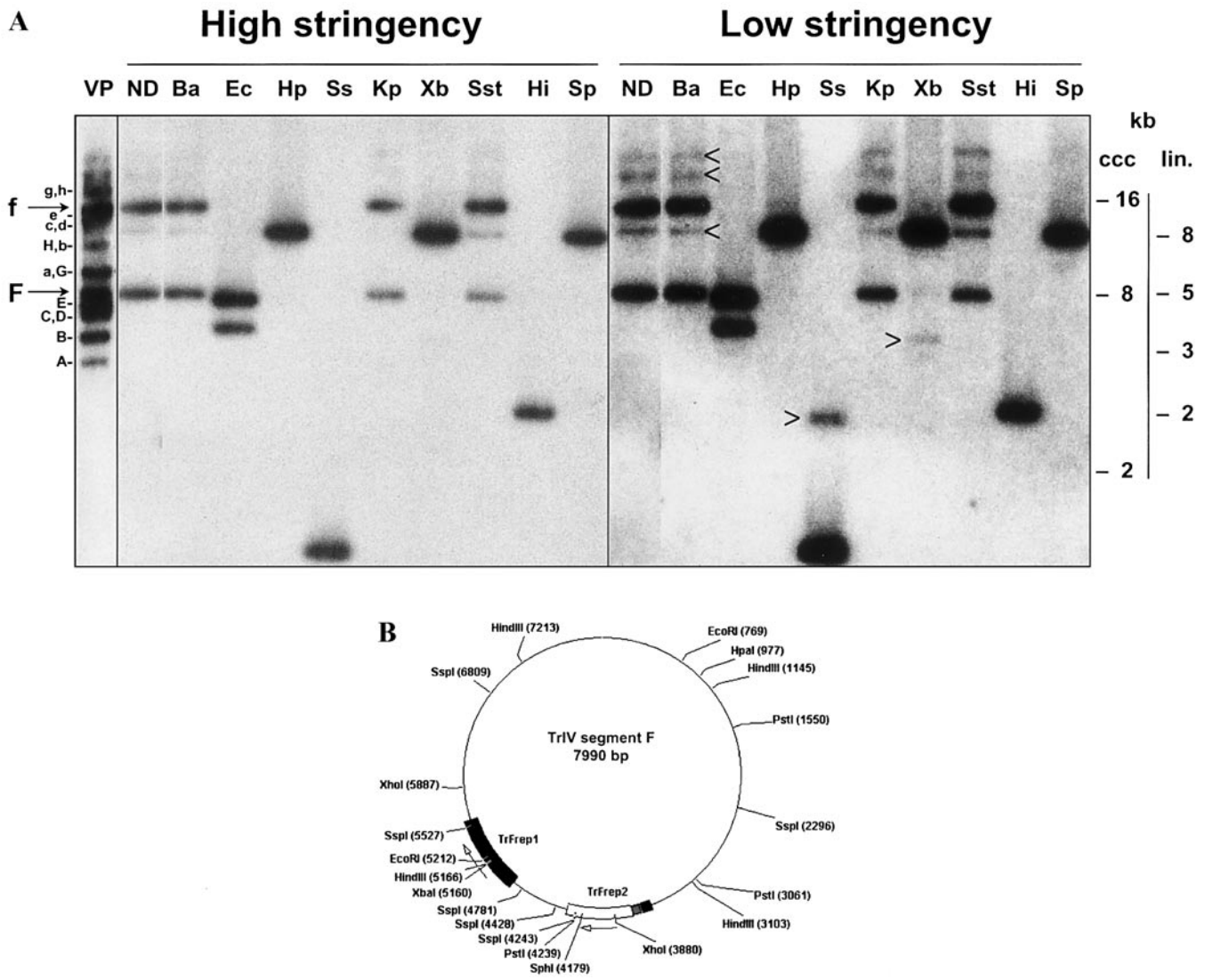


FIG. 3. TrIV genome segment F contains 2 potential *repeat element* genes. (A) Southern analysis of undigested and digested TrIV DNA using a *TrFrep1*-specific probe, under both high- and low-stringency conditions. In the left-most lane, undigested TrIV DNA was hybridized with ^{32}P-labeled viral DNA (same membrane as that used for hybridization with TrF*rep1*) to show the mobility of the various TrIV segments (compare with EtBr-stained gel in Cusson *et al.,* 1998). Abbreviations: VP, viral probe; ND, nondigested; Ba, *Bam*HI; Ec, *Eco*RI; Hp, *Hpa*I; Ss, *Ssp*I; Kp, *Kpn*I; Xb, *Xba*I; Sst, *Sst*I; Hi, *Hin*DIII; Sp, *Sph*I. Mobility of MW standards shown on the right is for covalently closed circular (ccc) and linear DNA. The TrIV genome segment that hybridized most strongly to the ^{32}P-labeled TrF*rep1* probe was identified as segment F, with the upper and lower bands seen here representing relaxed circular (f) and superhelical (F) species of the same segment. Left-pointing and right-pointing arrowheads: additional bands observed in undigested and digested TrIV DNA, respectively, under conditions of low stringency. (B) Restriction map of TrIV segment F showing the position of the TrF*rep1* and TrF*rep2* ORFs. In the latter, the white box represents the identified ORF, beginning with an ATG start codon, while the gray box encompasses additional nucleotides encoding amino acids displaying significant similarity to TrF*rep1,* but lacking an ATG start codon; further upstream (black box) lie a few additional nucleotides encoding amino acids unrelated to TrF*rep1,* bound by a TGA stop codon at the 5′ end. The TrF*rep1*-specific probe employed for the above Southern analysis is the *Ssp*I (4781–5527) fragment, encompassing most of the TrF*rep1* ORF.

All of the restriction enzymes that digested genome segment F (*Eco*RI, *Hpa*I, *Ssp*I, *Xba*I, *Hin*dIII, and *Sph*I) generated a labeling pattern consistent with the restriction map of this segment and with there being a single copy of TrF*rep1* in the TrIV genome (Figs. 3A and 3B). Under conditions of low stringency, the *Ssp*I digest generated an additional band (arrowhead) with a mobility corresponding to a fragment of TrF*rep2.* However, the mobility of the additional faint band observed in the *Xba*I digest (arrowhead) did not correspond to any predicted fragment.

The HdIV and TrIV *repeat element* genes are transcribed in the lepidopteran host

Patterns of expression of the ichnovirus *repeat element* genes in the infected lepidopteran host were analyzed by Northern blot or RT-PCR.

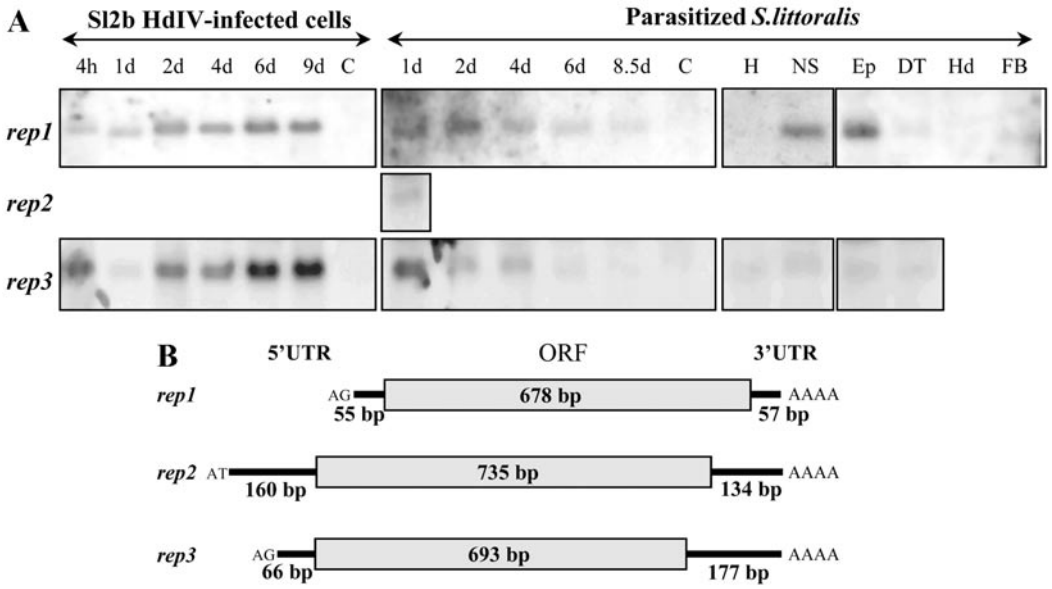


FIG. 4. Expression of HdIV genome segment E *repeat element* genes. (A) Northern analysis of total RNA extracted from Sl2b cells (~25 μg) and whole *Spodoptera littoralis* larvae (~50 μg) or specific tissues (~25 μg) at different times following infection with HdIV or parasitization by *Hyposoter didymator*. The same blot was used for hybridization with HdE*rep1* and HdE*rep3* probes. The *repeat element* specific probes were restriction fragments encompassing the corresponding genes, generated from the cloned segment E or the cloned C28 cDNA. For HdIV-infected Sl2b cell cultures and parasitized *Spodoptera littoralis* whole larvae, RNA was extracted at various times following treatment (4 h to 9 days; 8.5 days for parasitized larvae just before parasitoid pupation). *S. littoralis* larvae were parasitized by *H. didymator* 48 h after hatching; controls (C) were unparasitized, 4 days old, 2nd instar larvae. Some parasitized larvae were dissected 4 days postparasitization to isolate the hemolymph (H), the nervous system (NS), the epidermis (Ep) corresponding to cuticular epithelium and associated muscle, the digestive tract (DT), the fat body (FB), and the developing parasitoid larva (Hd). A same amount of RNA was loaded for each lane corresponding to a given sample. Based on a comparison of their mobility with that of RNA standards run on the same gel, size of all transcripts is in the 0.9–1.0 kb range, which is consistent with the predictions from RACE data. (B) Schematic representation of the three HdIV *repeat element* transcripts according to RACE data. Lengths of 5′ untranslated region (5′UTR), open reading frame (ORF), and 3′ untranslated region (3′UTR) are indicated for the three genes.

The HdIV *repeat element* genes were expressed in infected cell cultures as early as 4 h after infection, and until 9 days postinfection, as well as in larvae from 1 to 8.5 days postparasitism (Fig. 4A). Northern blot analysis indicated that HdE*rep1* and HdE*rep3* genes were abundantly transcribed in infected lepidopteran cells (Fig. 4A, *rep1* and *rep3* panels). On the other hand, HdE*rep2* transcripts were detected only in parasitized larvae and after several days of exposure (Fig. 4A, *rep2* panel), suggesting a level of expression lower than that observed for the other two genes. Moreover, because of the sequence similarities between the *repeat element* genes, detection of HdE*rep2* transcripts over prolonged exposure times could be due to cross-reacting hybridization. Indeed, cross-hybridization controls revealed a faint hybridization signal from both HdE*rep3* and HdE*rep1* using a HdE*rep2*-specific probe, but no cross hybridization between HdE*rep1* and HdE*rep3* fragments (data not shown). Because of the difficulty in detecting HdE*rep2* transcripts by Northern blot analysis, expression of the genes was verified by RACE experiments with RNAs extracted from HdIV-infected cell cultures. For all three genes, 5′RACE and 3′RACE products of expected size were obtained and verified by sequencing (Fig. 4B). Thus, HdE*rep1,* HdE*rep2,* and HdE*rep3* transcripts were 790, 1029, and 936 nt, respectively (Fig. 4B). These results confirmed expression of the HdE*rep2* gene in infected lepidopteran cells, though probably at a lower level than that observed for HdE*rep1* and HdE*rep3* genes.

Northern blot analysis indicated that HdE*rep1* and HdE*rep3* genes are expressed in parasitized *Spodoptera littoralis* larvae over the duration of parasitoid development (Fig. 4A). Transcripts were also detected in all tissues tested from parasitized larvae: hemocytes, nervous system, epidermis, digestive tube, and fat body (Fig. 4A, lanes H to DT and FB), but not from the parasitoid larvae (Fig. 4A, lane Hd). For both, a weak signal was detected in hemocytes. Interestingly, compared to the HdE*rep3* transcripts, the HdE*rep1* transcripts were more abundant in the nervous system (NS) and epidermis (Ep) than in the other tissues (Fig. 4A, compare lanes H, NS, Ep, and DT for HdE*rep1* and HdE*rep3*).

With respect to TrIV, a TrF*rep1*-specific transcript of ~900 bp was observed in all calyx fluid-injected and parasitized *C. fumiferana* 6th-instar larvae (Figs. 5A–5C). The transcript size estimated from the blot is that expected for TrF*rep1,* based on the length of the partial cDNA obtained (838 bp; part of 5′UTR missing) and on the predicted length of the 5′UTR (determined from an

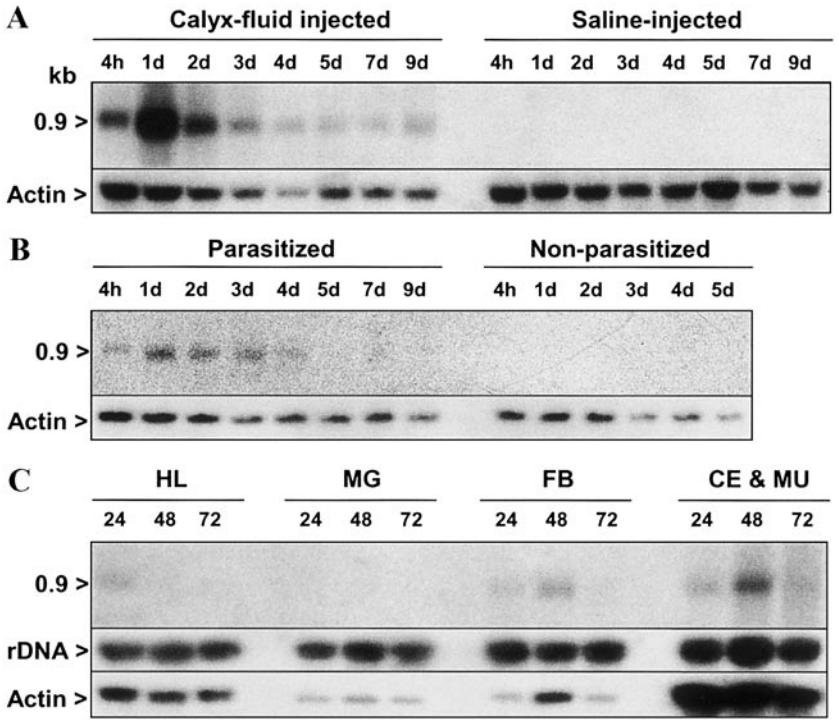


FIG. 5. Northern analysis of RNA (total) extracted from whole 6th-instar *Choristoneura fumiferana* larvae or specific tissues, at different times following parasitization by *Tranosema rostrale* or injection of its calyx fluid. The TrF*rep1*-specific probe was the same as that used for Southern analysis. (A) Larvae were injected with 0.5 FE of calyx fluid (in 1 μl saline) or 1 μl saline, 24 h after the molt to the 6th instar. At various times following treatment (4 h to 9 days), subgroups of larvae were chosen for RNA extraction. (B) Larvae were parasitized by *T. rostrale* 24 h after the molt to the 6th instar; control larvae were unparasitized animals of the same ages. Sampling for RNA extraction was conducted as in (A) except that no control larvae were available beyond the 5-day sampling point (they had all pupated). (C) Larvae were parasitized as in (B), and bled and dissected 24, 48, and 72 h postparasitization to isolate the hemolymph (HL), the midgut (MG), the fat body (FB), and the cuticular epithelium and associated muscle (CE & MU). The picture shown under each blot represents the same membrane that was rehybridized with a *Physarum polycephalum* ardB actin (Hamelin *et al.,* 1998) probe to monitor the amount of RNA present on the blot; in the case of (C), the membrane was also rehybridized with the *Coprinus cinereus* rDNA pCc1 clone (Wu *et al.,* 1983) as transcription of actin tends to vary from one tissue to another (note the much higher level of actin transcription in the CE & MU lanes than in the other lanes, likely due to the presence of muscle tissue). Size of transcripts based on a comparison of their mobility with that of RNA standards run on the same gel.

analysis of the promoter region). Unlike that which has been reported for *TrV1* (Béliveau *et al.,* 2000), parasitized and virus-injected larvae displayed very similar temporal patterns of TrF*rep1* transcription, with peak transcription observed 1 day after injection or parasitization (Figs. 5A and 5B). In a tissue-specific analysis, the bulk of TrF*rep1* transcription was observed in the fat body and the cuticular epithelium/muscle, with no signal detected in the midgut and only a weak signal detected in hemocytes 24 h postparasitization (Fig. 5C).

In an effort to determine whether TrF*rep2* represents a functional gene, mRNA extracted from TrIV-injected *C. fumiferana* larvae was reverse transcribed, followed by PCR amplification using a primer corresponding to a sequence common to both *repeat elements,* either in conjunction with the Abridged Universal Amplification Primer (containing a sequence present in all reverse transcription products) or with a primer specific to either one of the two *repeat elements.* Amplimers of the expected sizes were obtained with all three primer combinations. Digestion of the 710-bp amplimer (generated using primers common to both *repeat elements*) with TrF*rep1*-specific enzymes (*Eco*RI or *Hin*dIII) yielded the expected products, but a faint 710-bp signal remained detectable, possibly representing a transcript generated by TrF*rep2.* Similarly, digestion of the TrF*rep2*-specific 168-bp amplimer with *Xho*I (specific to TrF*rep2*) yielded the expected products (data not shown). These results indicate that a *repeat element* gene other than TrF*rep1* is transcribed in TrIV-infected *C. fumiferana* larvae, albeit at a lower level. However, whether this transcript is generated by TrF*rep2* or by a close homolog on another TrIV segment remains to be determined.

Sequence similarities among ichnovirus *repeat element* genes

The ClustalX alignment shown in Fig. 6 provides a comparison of the predicted proteins encoded by the *repeat element* genes from HdIV genome segment E, TrIV genome segment F, and CsIV genome segments B (CsB*rep1*) and I (CsI*rep1,* CsI*rep2,* and CsI*rep3;* Hilgarth

```
            1                                                                                                  90
                                                                                 ⇓                    * *         *
   HdErep1  MESKENNVLP AFSVLQGAPL QREIVIPLDI ILHLGDFLRF EDYRNFVKSI WPANDECDAV RNKLW.QRST HKIAIEFFNE EILNIEYNFD
   HdErep2  MNSSNNNVSR KTPAVPPLPA SMNVSLSIKK IFQAAESLEF EEYQSFIQKL LHNMDMHPQI QAQLW.RMST HRFTATFLNG KPLVIRYNYD
   HdErep3  ~~~~~~MTQY LS..TPNQST SEPVELPLRV IVYMARFLSV ADYRSFVKSI WSDEIPSKTV RTKLW.RMST RKITTKFING EPIEIEYNYD
   TrFrep1  ~~~~~~~MRI II..GFESIH ASSLLWPTEP SGRVRSFVQL RRIKRKRRTM MSPQNEMSPP DVCLPNVLNV KNFEATFYNG RRLDIQYDFN
   TrFrep2  ~~~~~~~~~~ ~~~~~~~~~H ASSLFGRVNV TGRVRFCVQL QRIKRKLRRM MSLENKISPP QVCLTNLLNV KDFEATFYNG RRLSIQYDFD
CsIrep1-R1  ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~MGSLTSAM LSSDEEAGTS EELLG.QQST RKFEATFLNG KEIEVEYTYE
CsIrep1-R2  ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~TP QEILQ.RLST HKFRATFLNG KEIEVEYIYH
   CsIrep2  MDSPVKEESR TL..TPAYFT SRQIFLPMEI IFHLSHFLEF QDFRNFIRAM WPNGEEDETF QERLR.QLST RKLVAAFLNG REIEVEYKYD
   CsIrep3  MESLGEGTSR TI..TPVYFT SRQVFLPTEI IFHLSHFLEF QDFRNFVRAM WPQWRRRRHF QRKTTGNFQH ERLLLLFLNG KTIEVEYNYD
   CsBrep1  MESLETKESK VF..TPVYFT SRQILLPLKM ISYVSRFLKF EDFRKFIRAM WPNGEANVIF QELLE.RLSI RKFKAKFYNR EEIEVEYKFD
    MOTIFS                                                                                 ST HKFEATFYNG KEIEIEYNYD

            91                                                                                                180
                          *   *           *                       *       *                                 *     **
   HdErep1  ASRTKDQQFL FNVETLSPVF GGVVPPGTN. QFLSASKLEN FLRMHVHLNM CSR.RQFAAC SCHELKCGTY TGVKIVKPPK VACRYGHFHH
   HdErep2  PSRLEEERVL FDVEYLFPVL GGVIPRALS. RFATASQIHS FVKKHVHLNR CAD.CEHASC PCDLGHD.QA QVRAFVQPAV DACHDRCFHH
   HdErep3  PGRIEEERVL INTKYLLPIS GGIVAPVPK. TFTSLTRINN FIQSAVELNV CSG.HEYACC PCHLKNFNRH SATAFAKPSA TACASKHFHH
   TrFrep1  PEKIVAERLR LNLESLLPLF GGIAPPGKA. EFTTIDEISN FVN.IVNLDS CSS.GIYASC SCHYNIP... EKEQ...NLV YRCRDDHFHH
   TrFrep2  RERIAEERLR VNLESLAPLF GGIVPPGEA. AFVTPSKIFN FVR.IIHMDN CKL.GIYASC SCRQNIM... KCEE...DSV DRCEDGHFHH
CsIrep1-R1  REGLDGIRIR INLQALLPVL GGIISPDTR. KFVSIFHVKD LLKTNPNLHE CSVYGLYNSC FNL..... DGGLE..EGG DKCPHDHFHH
CsIrep1-R2  REGLERLRIA VNLQGLLPVL GGFISTDAK. KFVSPFHVKD LLKTNPNLHE CSVYGLYNSC FNL..... DESLE..EEG DTCPYDHFHH
   CsIrep2  RERLLEDRVE VNLRSVLPIL GEVELPDGN. EFVNTLTMKE FLSKNVKLHK CEQ.GIHSFC RNVGVDP... DEYVDDYLEE DVCPQRHFHH
   CsIrep3  RERFLEDRVK INLQSMLPIL GEVELPDGN. EFVNTLTMKE FLYKNVKLHK CEQ.GIYSSC GVVAVDL... DE.....EER DVCPHHHFHH
   CsBrep1  RERSGINWIL INFKDLLPIL GGIMLPDDED KFQSIFTLED FLKRNLKVHR CSG.GIHTSC HNLGRDS... DSDSE..AKL DICPFDHYHH
    MOTIFS  RERLEEERIL INLESLLPIL GGIVPPDTN. EFVS                                                   CPCGHFHH

            181                                                                                          264
                *   *     *
   HdErep1  FCSQHVRDWV DIFLMSAVVK K........E EGSPSDADMT KRLLAYMRD. ......SVRL SGC~~~~~~~ ~~~~~~~~~~ ~~~~
   HdErep2  YCSQHVGYWL KLYLAPVVLL RERRASSADD RAAAESFLVF LSETVYFRG. .....LNVQL RDSPLQSVPS WKRR~~~~~~ ~~~~
   HdErep3  YCSNHVNHWF NSFLNSVIRS QEG.E.EPFN EDEAEIRLFL LDNMIYFRD. .....NEIKL RSSHLYRVL~ ~~~~~~~~~~ ~~~~
   TrFrep1  YCASHVSAWF KLYLERAILL QDESK.QFYY QLIADIHGPI TAFEFTVGR. ..YATAQYWL EYAR...... THVGRRRL~~ ~~~~
   TrFrep2  YCATHVESWL RFYLELAILL Q.ESD.DYYR QRMHQIYGRK LPCNVPLGG. ..PGTAEYWL KYARSYRLCS KHVTPKGSII LSI~
CsIrep1-R1  YCPEHVNAWF KYYLTVAIIL REG...DYAE LVKLSNPKDA DYLTTEIRS. ....NARYWL RRAQR.QRCR FSRE~~~~~~ ~~~~
CsIrep1-R2  YCPEHVNAWF KYYLTVAIIL REG...DYVE VVKLSNPQDA DYLTTEIRS. ....TAEYWL RYARR.KECG FSRKQLQSNY SV~~
   CsIrep2  YCPNHINRWF REYLSLVIAL QNGRQ.DFAD MIEFIYGEIG CEIAVNLRDI LTPLRLWLWF VGTR..KWCS VSD~~~~~~~ ~~~~
   CsIrep3  HRPRHINRWL REYLSPMIAL KNGQQ.DFVE MIEFTCGEFA HEIVANLWN. .SPLKLLQWF YELTNCSLCL DDD~~~~~~~ ~~~~
   CsBrep1  FCPDHVIAWF KHYLLTAILL REG...VYDE LVKNANLPNA DHLTSGRRR. ....TEQYWL RVARR.KKCR FSQ~~~~~~~ ~~~~
    MOTIFS  YCPQHVNNWF KKYLQWAILL QEG
```

FIG. 6. Alignment of *repeat element* encoded proteins. Clustal X alignment of the three HdIV (HdE*rep1,* HdE*rep2,* HdE*rep3*), the two TrIV (TrF*rep1* and TrF*rep2,* for the latter, only the sequence after the hypothetical intron is indicated), and CsIV (CsI*rep1* R1 & R2, CsI*rep2,* CsI*rep3,* and CsB*rep1*) predicted proteins. The beginning of the repeated element encoded amino acids is indicated by a vertical arrow; first amino acid is in bold. CsI*rep1* protein contains two copies of the single repeat sequence; the amino acids encoded by the repeats are divided into repeat 1 (R1) and repeat 2 (R2). The asterisk (*) indicates positions where absolute amino acid consensus occurs. MOTIFS indicates the consensus amino acid sequence identified as a motif by BLOCKMAKER.

and Webb, 2002a). As N- and C-termini represented more divergent regions, comparative analysis focused on the ~540-bp single repeat. Relative to the previously described CsB*rep1* protein, this region started at amino acid +59 for HdE*rep1* and HdE*rep2,* +51 for HdE*rep3,* +50 for TrF*rep1,* and +40 for the truncated TrF*rep2* (Fig. 6). The consensus for the single-repeat contains 145 amino acids (including gaps) and follows the pattern: x(17,18)-F-x(2)-N-x(8)-Y-x(20)-P-x(2)-G-x(8,9)-F-x(17,18)-C-x(7,8)-C-x(13,22)-C-(5)-H-H-x(4)-H-x(3)-W-x(4)-L-x(7,9).

Within this region, the three HdIV proteins were 25% identical, sharing 36 amino acids. In pairwise comparisons, amino acids identities were 37, 39, and 41% for HdE *rep1/rep2, rep1/rep3,* and *rep2/rep3,* respectively. If the entire protein was analyzed, the three HdIV proteins displayed 18% identity with pairwise identities ranging from 26 to 31%. For the TrIV proteins, amino acids identity inside the repeat was 66%, with the two predicted proteins sharing 89 amino acids. This high identity of the two TrIV proteins relative to HdIV or CsIV *repeat element* similarities suggests that the two TrIV genes have diverged recently. Pairwise comparisons between TrIV and HdIV proteins gave identities ranging from 28 to 32% within the single repeat.

In both HdIV and TrIV predicted proteins, four cysteine residues are conserved according to a C-x(7)-C-x(17,22)-C-x(8)-C motif (amino acids 141–181, Fig. 6), suggesting possible formation of disulfide bonds. The same region also contains a conserved C- and H-rich motif, C-x(3)-[HC]-[FY]-H-H-[FY]-C-x(2)-H-V (amino acids 172–185, Fig. 6) that suggests possible Zinc interactions. Finally, another region was rich in L, P, and G residues, following a pattern L-x(3,4)-[EK]-x-L-x-P-[LIV]-[FLS]-G(2)-[IV]-x-[PAL]-P (amino acids 100–116, Fig. 6). None of the three patterns generated significant matches in a scan of protein databases (SWISSProt and TrEMBL).

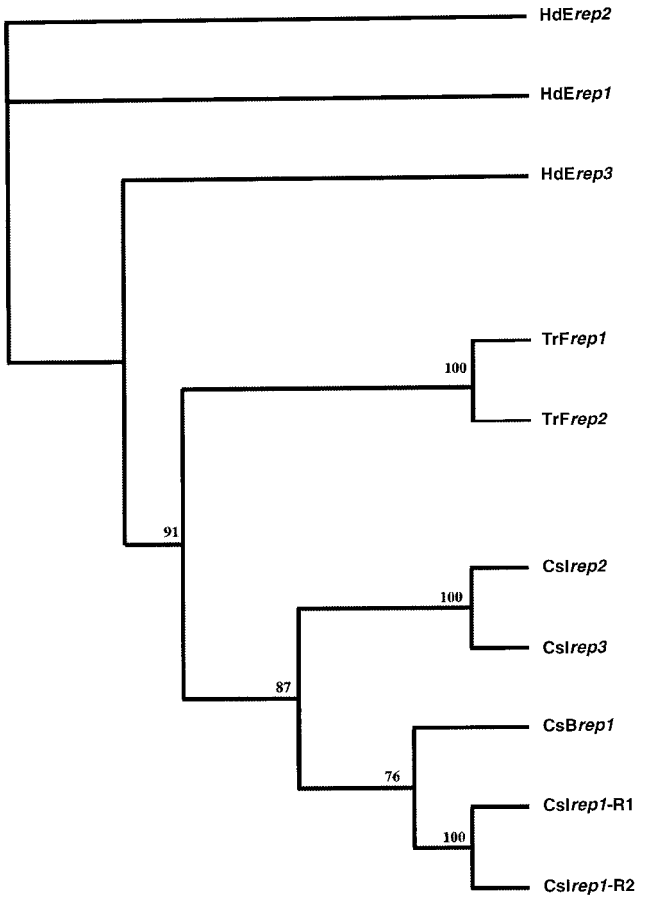


FIG. 7. Phylogenetic tree based on amino acid sequence of known Ichnovirus *repeat element* genes. Heuristic tree generated by PAUP based on the CLUSTAL X alignment of the single-repeat protein sequences with HdIV (HdE*rep1,* HdE*rep2,* and HdE*rep3*), TrIV (TrF*rep1* and TrF*rep2*), and CsIV (CsI*rep1,* CsI*rep2,* CsI*rep3,* and CsB*rep1*) predicted proteins (Fig. 6). Analysis predicted one tree with a total tree length of 910. Only Bootstrap values ≥ 50% are shown. CsI*rep1* protein is encoded by two tandem repeated sequences which were separated into repeat 1 (CsI*rep1*-R1) and repeat 2 (CsI*rep1*-R2).

All *repeat element* gene-encoded proteins lacked an N-terminal signal peptide and were predicted to be intracellular proteins. Indeed, preliminary experiments with transfected Sf9 cells allowed detection of a recombinant-HdE*rep1* protein in the cellular fraction using an anti-V5-tag antibody (unpublished observations). For all the proteins analyzed, searches in databases revealed no significant similarities with other known proteins or motifs.

Phylogenetic analysis of the *repeat element* gene family in ichnoviruses

A heuristic tree was generated based on the ClustalX alignment of the different ichnovirus *repeat element* encoded proteins (Fig. 7). The tree was drawn using the whole protein sequences, with the CsI*rep1* protein having its tandem repeated element separated in two repeats (R1 and R2). A similar analysis conducted with only the amino acids encoded by the repeat sequences resulted in an identical phylogenetic tree with slightly different bootstrap values (data not shown). Under bootstrap analysis, the relationships among the HdIV *repeat element* encoded proteins were indiscernible due to bootstrap values for the HdIV nodes being below threshold (≤50). As expected, the tree indicated that the three ichnovirus species are well separated; interestingly, TrIV genes appear to be more closely related to the CsIV genome segment I and B *repeat element* genes than to those of HdIV. Not surprisingly, the *repeat element* genes are much more similar within a species of wasp than between species. This is seen in the CsIV *repeat element* genes which display a much higher intra- than interspecific degree of similarity, with the exception of some of the predicted *repeat element* genes located on genome segment Z (Hilgarth and Webb, 2002b). These results suggest that the most significant duplication, divergence, and expansion of the *repeat element* genes occurred after speciation.

DISCUSSION

In the present article, we show that the *repeat element* gene family, originally described in CsIV, exists in two additional ichnoviruses, HdIV and TrIV. This is the first viral gene family shown to be conserved among different ichnovirus species, although early comparative hybridization analyses were indicative of variable degrees of genetic relatedness within polydnavirus genera (Stoltz and Whitfield, 1992). Although common in eukaryotes, gene families are rare among viruses. In baculoviruses, for example, some genes, including the *bro* genes and a few others, are organized in gene families, but most of them are unique (Hayakawa *et al.,* 2000). In contrast, polydnavirus gene families appear to be common, having been described in the ichnoviruses CsIV (Webb, 1998) and HdIV (Volkoff *et al.,* 1999), but also in the *Microplitis croceipes* bracovirus (Trudeau *et al.,* 2000). The CsIV ~540-bp repeat was originally reported as being ubiquitous in the genome (Theilmann and Summers, 1987). These sequences are now known to constitute a gene family of 28 members (Hilgarth and Webb, 2002a) that contain the ~540-bp sequence repeated singly or in tandem arrays of up to five repeats (Hilgarth and Webb, 2002b). Here we document three *repeat element* genes in HdIV encoded by genome segment E. In TrIV, one *repeat element* gene, and a possible second one, have been identified on genome segment F. Although many fewer *repeat element* genes have so far been identified in HdIV and TrIV than in CsIV, Southern blot analysis revealed the existence of additional HdIV and TrIV segments cross-hybridizing with *repeat element* probes, suggesting the existence of additional *repeat element* genes in these ichnovirus genomes.

Interestingly, polymorphism of genome segment E in HdIV was found to yield two additional slightly different *repeat element* genes. This variant of segment E, referred to here as segment E_{Var}, shares two-thirds of its sequence with segment E and is present at lower molarity in the virus population. Naturally occurring genetic variability exists within polydnavirus populations without apparent harm to the wasp (Fleming, 1991). Previously reported examples of polydnavirus genome polymorphism consisted mainly of differences in electrophoretic profiles of segments or restriction fragment length polymorphism (Stoltz and Xu, 1990). For example, in the *Hyposoter fugitivus* Ichnovirus, two variants of genome segment K were also detected but only differing by one restriction site (Stoltz and Xu, 1990), while the two variants of HdIV segment E differ by >1 kb. We have shown that these two HdIV genome segment E variants are present in the circular form of the virus in all individuals since both were amplified from the replicative tissue of individual females (Fig. 2B, lanes 5–6). Thus, segments E and E_{Var} are probably two distinct genome segments as both present distinct regions of integration. However, since the mechanisms of PDV replication are still poorly understood, segment E_{Var} could derive from segment E during one of the replication steps. Polydnaviruses have been proposed to be amplified by a rolling circle mechanism used for replication by some circular DNA viruses, such as Herpes virus (Boehmer and Lehman, 1997). In HdIV segment E, there is a direct 103-nt repeat on each side of the variable region. These repeats are 78% identical, with the last 16 nt perfectly conserved. The link between this repeated sequence, the genome segment replication, and the variable region should be further investigated.

With respect to gene expression, TrF*rep1* and the three *repeat element* genes on segment E of HdIV were shown to be transcribed in lepidopteran cells. Expression of HdIV segment E_{Var} *repeat element* genes was not investigated due to the lack of specific probes. Interestingly, differences in transcription levels were observed. As reported for the three CsIV *repeat element* genes encoded by genome segment I (Hilgarth and Webb, 2002a), the HdE*rep2* transcripts could only be detected by RT-PCR. In HdIV, disparities in levels of transcription among the *repeat element* genes may result from differences in promoter or regulatory regions since gene dosage regulation cannot explain such a difference. Interestingly, no consensus for TATA boxes were found on regions upstream the HdIV *repeat element* genes, although most of the known polydnavirus genes do have this element (Cui and Webb, 1997b; Dib-Hajj *et al.,* 1993), including TrF*rep1.* In the parasitized host, polydnaviral genes are rapidly expressed to varying degrees but usually show little temporal or tissue-specific regulation of transcription. However, exceptions have been noted. For example, temporal patterns of TrF*rep1* transcription in parasitized 6th-instar *C. fumiferana* larvae (Fig. 5; peak transcription 1 day postparasitism) were different from those of *TrV1* (Béliveau *et al.,* 2000; peak transcription 4 day postparasitism; same RNA used for both analyses), suggesting differences in the regulation of transcription of these two genes. Similarly, the three genes identified from *Chelonus inanitus* Bracovirus displayed different patterns of transcription, suggesting some regulation by host factors (Johner *et al.,* 1999).

Tissue-specific regulation of polydnaviral gene expression in the lepidopteran host remains a largely unexplored area. In general, hemocytes appear to be the predominant tissue expressing polydnavirus genes in the lepidopteran host, even if cells in other tissues, particularly the fat body, are also infected. TrF*rep1* showed a different pattern of tissue specificity in its expression since it is little expressed in hemocytes but abundantly expressed in the epidermis/muscle, as previously shown for *TrV1* (Béliveau *et al.,* 2000). On the other hand, HdE*rep1* seems to be expressed to a greater degree in nervous tissues and epidermis/muscle, which is not the case for HdE*rep3.* These examples illustrate the fact that the possible involvement of host factors in the regulation of polydnavirus gene transcription is still poorly understood.

Thus, for each of the studied ichnoviruses, several *repeat element* proteins appear to be expressed at different levels in their lepidopteran host. These proteins are predicted to be intracellular but their function is still unknown. No homology was found between *repeat element* genes and known proteins. However, Schmid and Tautz (1997) note that some functional protein modules are yet to be discovered or may evolve so rapidly that their homologs cannot be identified over larger evolutionary distances. In the lepidopteran host larva, polydnaviruses are believed to disrupt the immune system, inhibit host growth and development (Dover *et al.,* 1988), and alter host protein titers (Beckage *et al.,* 1987), in particular storage proteins (Béliveau and Cusson, 2001; Shelby and Webb, 1997). Since the *repeat element* genes reported here appear to be little expressed in hemocytes, we do not expect them to play a direct role in immune suppression. Even though the putative proteins described here are homologous, they may not all have the same function. Indeed, similar proteins may possess diversified functions in different organisms as some homologous transcription factors are known to have distinct functional roles (McClintock *et al.,* 2001; Wegner and Riethmacher, 2001). Organization of the *repeat element* genes in gene families increases the number of functionally active genes, as this organization provides an efficient mechanism for generating genetic diversity (Ohta, 2000). For example, animal toxins are often encoded by multigene families and appear to have diversified by gene duplication and adaptative evolution (Kordis and Gubensek, 2000). Conservation of a gene family

among at least three ichnoviruses therefore suggests an important biological function for these *repeat element* genes. Given that the wasp hosts of the three ichnoviruses examined here can each successfully parasitize more than one lepidopteran host species, the possibility that *repeat element* variants are required for completion of wasp development in these different host species cannot be discounted. A comparison of the relative transcription levels of *repeat element* genes (of a given ichnovirus) in different lepidopteran hosts should help establish whether such a mechanism is indeed plausible.

Thus far, *repeat element* genes have been identified from ichnoviruses associated with the ichneumonid subfamily *Campopleginae.* That an ancestor gene originated from one of the two hosts (hymenoptera or lepidoptera) seems plausible since horizontal gene transfer, a major evolutionary force in bacteria, is also known to occur in eukaryotes (Campbell, 2000). In addition, viruses are known to have evolved by capturing and using genes from infected cells (Becker, 2000). When sequences become available for host and viral genomes, comparative studies may help in determining from which host, if any, *repeat element* genes originated. Orthologs and paralogs probably arose then from gene or gene cluster duplication events. Indeed, more than a third of a typical eukaryotic genome consists of duplicate genes and gene families (Wagner, 2001). Because of the sequence similarities among *repeat element* genes, intrasegment and/or intersegment recombinational events may have contributed to increase the number of genes within a species. Indeed, genome segmentation confers unique opportunities for virus recombination through exchange of genome segments, although the polydnavirus life cycle would seem to limit the potential for this type of segment exchange (Webb, 1998).

Conservation among ichnoviruses of the *repeat element* motif suggests that important functional domains of the proteins are likely to be conserved. Their diversity, conservation, and expression could indicate functional involvement in multiple physiological processes. Targeted analyses of repeat element proteins singly and in combination are now necessary to study the function(s) of these proteins.

MATERIALS AND METHODS

Biological material

Rearing of *S. littoralis* larvae and *H. didymator* wasps, handling of cell cultures, HdIV extraction, and infection of Sf9 and Sl2b cells were conducted as described in Volkoff *et al.* (1999).

C. fumiferana larvae and *T. rostrale* wasps were obtained as described previously (Doucet and Cusson, 1996a,b; Cusson *et al.,* 1998). Parasitization and calyx-fluid injection of 1-day-old 6th instar *C. fumiferana* larvae were carried out as described in Béliveau *et al.* (2000) and Doucet and Cusson (1996a,b), respectively.

Identification of *repeat element*-like sequences in HdIV and TrIV

PCR amplification and sequencing of the HdIV genome segment containing repeat element *like gene(s).* In an earlier study, we constructed a cDNA library from RNAs extracted from HdIV-infected Sf9 cells, 24 h postinfection (Volkoff *et al.,* 1999). The clones were sequenced using the Abi Prism Dye Terminator Cycle Sequencing Ready Reaction Kit (Perkin–Elmer).

To amplify the HdIV genome segment(s) containing the gene corresponding to the cDNA named C28, two specific primers were designed: **HdCS3** 5′-ATCCGACGCTGACATGACCAAG-3′, and **HdCS5** 5′-GCTGCGATATCAACTAACGACG-3′. Touchdown PCR amplifications were conducted with Taq DNA Polymerase (Gibco) and HdIV DNA as template (100–200 ng), using five cycles of 95°C, 30 s; 63°C, 30 s; 72°C, 3 min, followed by 25 cycles of 95°C, 30 s; 60°C, 30 s; 72°C, 3 min, and a final 5 min elongation step at 72°C. Amplification products were recovered from low melting temperature agarose gel using glass beads (Ausubel *et al.,* 1995) and cloned in 3′-T overhangs pGEM-TEasy vector system (Promega), according to the manufacturer's protocol (TM042). The inserts were sequenced using the Big Dye Terminator Cycle Ready Reaction DNA Sequencing Kit (Applied Biosystems, AbiPrism, Perkin–Elmer) in an ABI PRISM 310 Genetic Analyser (Perkin–Elmer). Overlapping nucleotide sequence fragments were assembled using Fragment Assembly, based on the method of Staden (1980), part of the Wisconsin Sequence Analysis Package (Version 8 for Unix Server, Genetics Computer Group).

Once these amplification products were sequenced, segment-specific reverse PCR primers were designed: for segment E, **F–E** 5′-ATCGGGTGCGGGTAGGGTTTGAG-3′ and **R–E** 5′-GACGAGGGAACACTCTGGAGTGGAGAA-3′; for segment E_{Var}, **F–E_{Var}** 5′-ACGCCAGTCTATTGTATCCTGTGCTA-3′ and **R–E_{Var}** 5′-GTTACCGTCGTCCACCAGTTTGAG-3′ (primer positions are indicated in Fig. 1). PCR was conducted using the Advantage 2 PCR Enzyme System (Clontech) according to manufacturer's protocol (PT3281-1). The cycling parameters consisted of 30 cycles of 95°C, 30 s; 68°C, 3 min, and a final 5-min elongation at 68°C. Cloning and sequencing were performed as above.

Isolation and sequencing of a TrIV clone containing repeat element-*like sequence.* In an earlier study (Béliveau *et al.,* 2000), a screen of two TrIV genomic libraries, using TrIV-specific cDNAs as a probe, generated several positive clones. One of them, an 8-kb *Sph*I clone (containing the Tr*Frep1* gene, here described as a *repeat element* homolog), was selected, subjected to restriction

enzyme mapping, and partially sequenced (DNA sequencing service, Université Laval, Québec, Canada).

Southern analysis

Restriction map analysis of HdIV genome segment E and identification of chromosomal integration region. DNA extractions and Southern blot hybridizations were conducted as described in Volkoff *et al.* (2001, 1999), respectively. Wasp DNA was extracted from AA-Hd and K-Hd cell lines established from wasp pupae (Rocher, unpublished observations). Genomic DNA (4 and 2 μg for wasps and HdIV, respectively) was digested with restriction enzymes having sites distributed on the genome segment E: *Bam*HI, *Bgl*II, *Eco*RI, *Eco*RV, *Hin*dIII, *Nsi*I, *Pst*I, *Sal*I. DNA was probed with the entire segment E (C28 cDNA and E amplimer, the latter corresponding to segment E, minus the HdE*rep1* gene region).

Identification of TrIV genome segment(s) encoding repeat element-like genes and assessment of copy number. TrIV DNA extraction and Southern blot hybridization were conducted as described previously (Béliveau *et al.,* 2000). The probe was a α^{32}P-labeled 747-bp *Ssp*I DNA fragment encompassing most of the coding region (from nt 4781 to 5527 of the sequence; Fig. 3B) of the TrF*rep1* gene. While hybridization and washes for the high-stringency detection were carried out as described earlier (Béliveau *et al.,* 2000), conditions for low-stringency detection were modified as follows: hybridization was carried out in 5× SSPE, 30% formamide, 5× Denhardt's reagent, and 200 μg/ml denatured salmon sperm DNA for 16 h at 43°C; three 10-min washes were performed at room temperature using 2× SSC and 0.1% SDS, followed by three other 30-min washes at 37°C in 0.1× SSC and 0.1% SDS. The blots were autoradiographed at −80°C for either 5 h (high stringency) or 16 h (low stringency).

PCR amplification of HdIV genome segment E related DNA from wasp DNA

DNA was isolated from individual *H. didymator* wasps (modified from Sambrook *et al.,* 1989), 50–100 ng (1/10) of which was used as template for PCR. Analyzed samples consisted of male wasps (n = 25), ovariectomized females, and corresponding reproductive tracts (n = 15). Amplifications were conducted with Taq DNA Polymerase (Gibco) using 5 cycles of 94°C, 30 s; 54°C, 30 s; 72°C, 3 min, followed by 25 cycles of 94°C, 30 s; 51°C, 30 s; 72°C, 3 min, and a final 5-min elongation at 72°C. The primers used were **F**$_1$ 5′-GTACTATCGGAGAAACAG-3′ (−180 nt before HdE*rep2* ORF), **R**$_1$ 5′-GACTTGACGAAGCTGCGA-3′ (+122 nt after ATG HdE*rep3*), and **R**$_2$ 5′-GATTCCGACGTCGATTGA-3′ (+319 nt after HdE*rep3* ORF). All primers were common to segment E and its variant segment E$_{Var}$. The relative positions of the primers are shown in Fig. 2C.

Northern analysis

Detection of HdIV repeat element *RNAs.* Northern blot hybridizations were conducted as described in Volkoff *et al.* (1999). Total RNAs were extracted from uninfected and HdIV-infected Sl2b and Sf9 cells, and from control and parasitized *S. littoralis* larvae, using the TRIZOL Reagent (Gibco-BRL) according to the manufacturer's instructions. Infected samples were collected 4 h, 1 day, 2 days, 4 days, 6 days, and 9 days after infection depending on the assay. Total RNAs were also extracted from tissue samples collected from parasitized *S. littoralis* 4d postparasitism: hemocytes, nervous tissue (nerve cord and head), digestive tract, epidermis (cuticular epithelium/muscle/tracheole), digestive tract (epithelium/muscle/tracheole), and fat body. Total RNAs from 3rd instar parasitoid larvae 3–4 days old were also collected. The *repeat element* specific probes used were pC28 *Eco*RI fragment for HdE*rep1,* pSHE *Sna*BI fragment for HdE*rep2,* and pSHE *Sca*I fragment for HdE*rep3.* Southern blot analyses were performed to determine if the probes cross-hybridized: 50 ng of each of the three probes was hybridized with HdE*rep1* and with HdE*rep2* probes.

Detection of TrIV repeat element *RNAs.* The procedure employed for Northern blot hybridization was the same as that described in Béliveau *et al.* (2000), using the same TrF*rep1* 747-bp *Ssp*I DNA fragment as probe as for the Southern analysis. Following hybridization, the membranes were stripped and rehybridized with the a *Physarum polycephalum* ardB actin (Hamelin *et al.,* 1998) probe to monitor the amount of RNA in each lane; for the tissue-specific analysis, the membrane was also rehybridized with a *Coprinus cinereus* pCc1 rDNA clone (Wu *et al.,* 1983) since actin transcription tends to vary among different tissues.

RT-PCR

Detection of HdIV repeat element *mRNAs: 3′RACE.* The cDNA template was synthesized from 6 μg of total RNA extracted as described above from Sl2b cells, either uninfected or 48 h p.i. The mRNA was primed with 500 ng of Q_T primer (C-Q_2-GAGCTCAAGC-$(T)_{17}$; Frohman, 1994) for 2 min at 70°C; reverse transcription reaction was conducted with Superscript RnaseH-Reverse Transcriptase (Gibco-BRL, Life Technologies) for 50 min at 42°C in an oven. One 200th of the resulting cDNA pool was used for PCR amplification. Amplification of the 3′ partial cDNA ends was conducted with the reverse primer **Q**$_2$ 5′-CAGTGAGCACAGTGACGAGGACTC-3′ and forward gene-specific primers: **3′RACErep1** 5′-GAGAACAATGTTTTGCCTGCCTTTTC-3′ (+13 nt after HdE*rep1* ATG), **3′RACErep2** 5′-GCTTCCAGCGTCAATGAATGTCAG-3′ (+50 nt after HdE*rep2* ATG), **3′RACErep3** 5′-ACCGGACGAGTTACCGTTGCGAG-3′ (+42 nt after HdE*rep3* ATG). PCR amplifications were conducted with Taq DNA Poly-

merase (Gibco) using five cycles of 94°C, 30 s; 65°C, 30 s; 72°C, 1′30 min, followed by 25 cycles of 94°C, 30 s; 62°C, 30 s; 72°C, 1′30 min, and a final 5-min elongation at 72°C. To control that amplification products did not arise from DNA contamination, PCR was also conducted with RNA template in an amount comparable to that of the cDNA. Cloning and sequencing of the 3′RACE products were done as described above.

5′RACE. The cDNA template for RT-PCR was synthesized from 3 μg of poly(A) mRNA, extracted with PolyATract Isolation System (Promega) from total RNA extracted as described above from Sf9 cells, either uninfected or 24 h p.i. The 5′ amplification of cDNA ends was performed with the SMART RACE cDNA Amplification Kit (Clontech), according to the manufacturer's instructions (PT3269-1). For amplification, the following gene-specific primers were designed: **5′RACErep1** 5′-GGGCTTCCCTCTTCCTTCTTCACCA-3′ (+620 nt after HdE*rep1* ATG), **5′RACErep2** 5′-CACGAGGGAACACTCTGGAGTGGAG-3′ (+722 nt after HdE*rep2* ATG), **5′RACErep3** 5′-CCTTATTTCGGCCTCATCCTCGTTG-3′ (+620 nt after HdE*rep3* ATG). 5′RACE products were cloned in pGEM-TEasy vector (Promega) and sequenced (GenomeExpress, Grenoble, France).

*Cloning of TrF*rep1 *cDNA.* One microgram of total RNA (from calyx-fluid-injected larvae, 2 days postinjection) was reverse transcribed using an Adapter Primer 5′-GGCCACGCGTCGACTAGTAC$(T)_{17}$-3′ (Invitrogen Life Technologies) containing a *Spe*I restriction site. The reaction was carried out at 42°C for 1 h. One-sixth of the reverse transcription reaction was used for PCR amplification (Béliveau *et al.,* 2000) with a TrF*rep1* specific primer (from nt +4881 to +4896, Fig. 3) containing a *Bam*HI restriction site at its 5′ end (5′-TCGCGGATCCATGCGCATTATCATC-3′) and the Abridged Universal Amplification Primer 5′-GGCCACGCGTCGACTAGTAC-3′ (Invitrogen Life Technologies) corresponding to the 5′ portion of the Adapter Primer. The PCR product was digested with *Bam*HI and *Spe*I and cloned in the corresponding sites of the pLitmus-29 vector (New England Biolabs) for sequence analysis.

RT-PCR detection of TrIV repeat element *mRNAs.* Following a reverse transcription step identical to that described above, a primer (P26) corresponding to a sequence common to TrF*rep1* (from nt +5086 to +5103) and TrF*rep2* (from nt +3805 to +3821) was used together with the Abridged Universal Amplification Primer to generate a 710-bp amplicon, which was then digested with enzymes known to cut only in TrF*rep1* (*Eco*RI and *Hin*dIII). Primer P26 was also used in conjunction with primer P27, specific to TrF*rep1* (complementary to nt +5416 to +5397), as well as with primer P28, specific to TrF*rep2* (complementary to nt +3973 to +3954); the two pairs were expected to generate fragments of 330 and 168 bp, respectively. The latter product was digested with an enzyme known to cut only in TrF*rep2* (*Xho*I).

Nucleotide and amino acid sequence analysis

ORFs were predicted using either ORF Finder at the NCBI website (http://www.ncbi.nlm.nih.gov) or GENSCAN (Burge and Karlin, 1997) at the MIT website (http://genes.mit.edu/GENSCAN.html). Comparisons of the ORF sequences with sequences in nonredundant databases (GenBank CDS translations + PDB + SwissProt + PIR + PRF) was carried out using BLAST (Altschul *et al.,* 1997) Sequence Similarity Searches available at different websites. Motifs were searched using GeneFIND family identification system at Protein Information Resource website (http://pir.georgetown.edu/gfserver/genefind.html). The SWISS-PROT and TrEMBL databases were scanned for the occurrence of *repeat element* amino acids patterns using ScanProsite program at ExPasy website (http://www.expasy.ch/tools/scnpsit2.html).

Multiple alignment comparisons were carried out using the ClustalW (Thompson *et al.,* 1994) Service at the European Bioinformatics Institute website (http://www2.ebi.ac.uk/clustalw). Final alignments were adjusted by hand to ensure that both amino acids and nucleotides were aligned.

Cellular localization of viral gene products was predicted using PSORT (Prediction of Protein Sorting Signals; Horton and Nakai, 1997) at NIBB website (http://psort.nibb.ac.jp).

Pile-up of the HdIV and TrIV *repeat element* genes with CsIV *repeat element* genes

Alignments of the HdIV and TrIV *repeat element* proteins were made to four identified CsIV *repeat element* proteins (SwissProt Accession No. P17578 for CsB*rep1;* Hilgarth and Webb, 2002a for the CsI*rep* genes). They were done using CLUSTAL X (Thompson *et al.,* 1997), which was then scanned to ensure optimal alignments. Amino acid domains encoded by the repeat portion of the *repeat element* genes were identified by using Blockmaker (Henikoff *et al.,* 1995). Phylogenetic analysis of the predicted *repeat element* proteins was performed using the PAUP 3.1 program (Swofford, 1993). The integrity of the trees produced by PAUP were checked by bootstrap analysis according to Felsenstein (1985).

ACKNOWLEDGMENTS

The authors thank B. Limier for providing *Spodoptera* larvae and S. Berger for help in reagent preparation. We also thank D. B. Stoltz for constructive comments on the manuscript. This work was partially supported by a USDA:NRI grant to B.A.W. (9802588) and is referenced as publication no. 02-08-005 of the University of Kentucky Agricultural Experiment Station. Work on TrIV was funded in part by the Canadian Biotechnology Strategy Fund, the Natural Science and Engineering Research Council of Canada (postgraduate scholarship to A.L.), and the Canadian Forest Service.

REFERENCES

Altschul, S. F., Madden, T. L., Schaffer, A. A., Zhang, J., Zhang, Z., Miller, W., and Lipman, D. J. (1997). Gapped BLAST and PSI-BLAST: a new

generation of protein database search programs. *Nucleic Acids Res.* **25,** 3389–3402.

Ausubel, F. M., Brent, R., Kingston, R. E., Moore, D. D., Seidman, J. G., Smith, J. A., and Struhl, K., Eds. (1995). "Current Protocols in Molecular Biology." John Wiley & Sons Inc., New York.

Beckage, N. E. (1998). Modulation of immune response to parasitoids by polydnaviruses. *Parasitology* Suppl.:S57–S64.

Beckage, N. E., Templeton, T. J., Nielsen, B. D., Cook, D. I., and Stoltz, D. B. (1987). Parasitism-induced hemolymph polypeptide in *Manduca sexta* (L.) larvae parasitized by braconid wasp *Cotesia congregata* (Say). *Insect Biochem.* **17,** 439–455.

Becker, Y. (2000). Evolution of viruses by acquisition of cellular RNA or DNA nucleotide sequences and genes: an introduction. *Virus Genes* **21,** 7–12.

Béliveau, C., and Cusson, M. (2001). Disruption of lepidopteran metamorphosis by polydnaviruses. *In* "Perspective in Comparative Endocrinology: Unity and Diversity" (H. J. Th. Goos, R. K. Rastogi, H. Vaudry, and R. Pierantoni, Eds.), pp. 165–172. Monduzzi Editore, Bologna, Italy.

Béliveau, C., Laforge, M., Cusson, M., and Bellemare, G. (2000). Expression of a *Tranosema rostrale* polydnavirus gene in the spruce budworm, *Choristoneura fumiferana. J. Gen. Virol.* **81,** 1871–1880.

Boehmer, P. E., and Lehman, I. R. (1997). Herpes simplex virus DNA replication. *Annu. Rev. Biochem.* **66,** 347–384.

Burge, C., and Karlin, S. (1997). Prediction of complete gene structures in human genomic DNA. *J. Mol. Biol.* **268,** 78–94.

Campbell, A. M. (2000). Lateral gene transfer in prokaryotes. *Theor. Popul. Biol.* **57,** 71–77.

Cui, L., Soldevilla, A. I., and Webb, B. A. (2000). Relationships between polydnavirus gene expression and host range of the parasitoid wasp *Campoletis sonorensis. J. Insect Physiol.* **46,** 1397–1407.

Cui, L., and Webb, B. A. (1997). Promoter analysis of a cysteine-rich *Campoletis sonorensis* polydnavirus gene. *J. Gen. Virol.* **78,** 1807–1817.

Cusson, M., Béliveau, C., Laforge, M., Bellemare, G., Levasseur, A., and Stoltz, D. (2001). Hormonal alterations and molecular mechanisms underlying the induction of host developmental arrest by endoparasitic wasps. *In* "Endocrine Interactions of Parasites and Pathogens with Their Insect Hosts" (J. P. Edwards and R. J. Weaver, Eds.), pp. 111–121. BIOS Scientific Publishers, Oxford.

Cusson, M., Laforge, M., Miller, D., Cloutier, C., and Stoltz, D. (2000). Functional significance of parasitism-induced suppression of juvenile hormone esterase activity in developmentally delayed *Choristoneura fumiferana* larvae. *Gen. Comp. Endocrinol.* **117,** 343–354.

Cusson, M., Lucarotti, C., Stoltz, D., Krell, P., and Doucet, D. (1998). A polydnavirus from the spruce budworm parasitoid, *Tranosema rostrale* (Ichneumonidae). *J. Invert. Pathol.* **72,** 50–56.

Dib-Hajj, S. D., Webb, B. A., and Summers, M. D. (1993). Structure and evolutionary implications of a "cysteine-rich" *Campoletis sonorensis* polydnavirus gene family. *Proc. Natl. Acad. Sci. USA* **90,** 3765–3769.

Doucet, D., and Cusson, M. (1996a). Role of calyx fluid in alterations of immunity in *Choristoneura fumiferana* larvae parasitized by *Tranosema rostrale. Comp. Biochem. Physiol. [A]* **114,** 311–317.

Doucet, D., and Cusson, M. (1996b). Alteration of developmental rate and growth of *Choristoneura fumiferana* parasitized by *Tranosema rostrale:* Role of the calyx fluid. *Entomol. Exp. Appl.* **81,** 21–30.

Dover, B. A., Davies, D. H., and Vinson, S. B. (1988). Degeneration of last instar *Heliothis virescens* prothoracic glands by *Campoletis sonorensis* polydnavirus. *J. Inv. Pathol.* **51,** 80–91.

Felsenstein, J. (1985). Confidence limits on phylogenies: An approach using the bootstrap. *Evolution* **39,** 783–791.

Fleming, J. G. (1991). The integration of polydnavirus genomes in parasitoid genomes: Implications for biocontrol and genetic analyses of parasitoid wasps. *Biocontrol* **1,** 127–135.

Fleming, J. G., and Summers, M. D. (1991). Polydnavirus DNA is integrated in the DNA of its parasitoid wasp host. *Proc. Natl. Acad. Sci. USA* **88,** 9770–9774.

Frohman, M. A. (1994). Cloning PCR products. *In* "The Polymerase Chain Reaction" (K. B. Mullis, F. Ferré, and R. A. Gibbs, Eds.), pp. 14–37. Birkhauser, Boston.

Gruber, A., Stettler, P., Heiniger, P., Schumperli, D., and Lanzrein, B. (1996). Polydnavirus of the braconid wasp *Chelonus inanitus* is integrated in the wasp's genome and excised only in later pupal and adult stages of the female. *J. Gen. Virol.* **77,** 2873–2879.

Hamelin, M., Adam, L., Lemieux, G., and Pallota, D. (1988). Expression of the three unlinked isocoding actin genes of *Physarum polycephalum. DNA* **7,** 317–328.

Hayakawa, T., Rohrmann, G. F., and Hashimoto, Y. (2000). Patterns of genome organization and content in Lepidopteran Baculoviruses. *Virology* **278,** 1–12.

Henikoff, S., Henikoff, J. G., Alford, W. J., and Pietrokovski, S. (1995). Automated construction and graphical presentation of protein blocks from unaligned sequences. *Gene-Combis* **163,** GC17–GC26.

Hilgarth, R. S., and Webb, B. A. (2002a). Validation of gene family status of the *repeat element* gene family in *Campoletis sonorensis* ichnovirus. Submitted for publication.

Hilgarth, R. S., and Webb, B. A. (2002b). Prevalence of the *repeat element* gene family in the *Campoletis sonorensis* ichnovirus. In preparation.

Horton, P., and Nakai, K. (1997). Better prediction of protein cellular localization sites with the *k* nearest neighbor classifier. *Intellig. Syst. Mol. Biol.* **5,** 147–152.

Johner, A., Stettler, P., Gruber, A., and Lanzrein, B. (1999). Presence of polydnavirus transcripts in an egg-larval parasitoid and its lepidopterous host. *J. Gen. Virol.* **80,** 1847–1854.

Kordis, D., and Gubensek, F. (2000). Adaptative evolution of animal toxin multigene families. *Gene* **261,** 43–52.

Kozak, M. (1989). The scanning model for translation: An update. *J. Cell Biol.* **108,** 229–241.

McClintock, J. M., Carlson, R., Mann, D. M., and Prince, V. E. (2001). Consequences of *Hox* gene duplication in the vertebrates: An investigation of the zebrafish *Hox* paralogue group 1 genes. *Development* **128,** 2471–84.

Ohta, T. (2000). Evolution of gene families. *Gene* **259,** 45–52.

Sambrook, J., Fritsch, E. F., and Maniatis, T. (1989). "Molecular Cloning: A Laboratory Manual." 2nd ed., Cold Spring Harbor Laboratory, Cold Spring Harbor, New York.

Savary, S., Beckage, N. E., Tan, F., Periquet, G., and Drezen, J. M. (1997). Excision of the polydnavirus chromosomal integrated EP1 sequence of the parasitoid *Cotesia congregata* (*Braconidae, Microgastrinae*) at potential recombinase binding sites. *J. Gen. Virol.* **78,** 3125–3134.

Schmid, K. J., and Tautz, D. (1997). A screen for fast evolving genes from *Drosophila. Proc. Natl. Acad. Sci. USA* **94,** 9746–9750.

Shelby, K. S., Cui, L., and Webb, B. A. (1998). Polydnavirus mediated inhibition of lyzozyme gene expression and the antibacterial response. *Insect Mol. Biol.* **7,** 265–272.

Shelby, K. S., and Webb, B. A. (1997). Polydnavirus infection inhibits translation of specific growth-associated proteins. *Insect Biochem. Mol. Biol.* **27,** 263–270.

Staden, R. (1980). A computer program to search for tRNA genes. *Nucleic Acids Res.* **8,** 817–825.

Stoltz, D. B., and Whitfield, J. B. (1992). Viruses and virus-like entities in the parasitic Hymenoptera. *J. Hym. Res.* **1,** 125–139.

Stoltz, D. B., and Xu, D. (1990). Polymorphism in polydnavirus genomes. *Can. J. Microbiol.* **36,** 538–543.

Swofford, D. L. (1993). Phylogenetic analysis using parsimony. Version 3.1. Illinois Natural History Survey, Champaign, IL.

Theilmann, D. A., and Summers, M. D. (1987). Physical analysis of the *Campoletis sonorensis* virus multipartite genome and identification of a family of tandemly repeated elements. *J. Virol.* **61,** 2589–2598.

Theilmann, D. A., and Summers, M. D. (1988). Identification and comparison of *Campoletis sonorensis* virus transcripts expressed from four genomic segments in the insect hosts *Campoletis sonorensis* and *Heliothis virescens. Virology* **167,** 329–341.

Thompson, J. D., Gibson, T. J., Plewniak, F., and Higgins, D. G. (1997). The Clustal X windows interface: Flexible strategies for multiple sequence alignment aided by quality analysis tool. *Nucleic Acids Res.* **25,** 4876–4882.

Thompson, J. D., Higgins, D. G., and Gibson, T. J. (1994). CLUSTAL W: Improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Res.* **22,** 4673–4680.

Trudeau, D., Witherell, R. A., and Strand, M. R. (2000). Characterization of two novel *Microplitis demolitor* polydnavirus mRNAs expressed in *Pseudoplusia includens* haemocytes. *J. Gen. Virol.* **81,** 3049–3058.

Varrichio, P., Falabella, P., Sordetti, R., Graziani, F., Malva, C., and Pennacchio, F. (1999). *Cardiochiles nigriceps* polydnavirus: molecular characterization and gene expression in parasitized *Heliothis virescens* larvae. *Insect Biochem. Mol. Biol.* **29,** 1087–1096.

Volkoff, A.-N., Cerutti, P., Rocher, J., Ohresser, M. C., Devauchelle, G., and Duonor-Cerutti, M. (1999). Related RNAs in lepidopteran cells after in vitro infection with *Hyposoter didymator* virus define a new polydnavirus gene family. *Virology* **263,** 349–363.

Volkoff, A.-N., Rocher, J., Cerutti, P., Ohresser, M. C., d'Aubenton-Carafa, Y., Devauchelle, G., and Duonor-Cerutti, M. (2001). Persistent expression of a newly characterized *Hyposoter didymator* polydnavirus gene in long-term infected lepidopteran cell lines. *J. Gen. Virol.* **82,** 963–969.

Wagner, A. (2001). Birth and death of duplicated genes in completely sequenced eukaryotes. *Trends Genet.* **17,** 237–239.

Webb, B. A. (1998). Polydnavirus biology, genome structure, and evolution. *In* "The Insect Viruses" (L. K. Miller and L. A. Balls, Eds.), pp. 105–139. Plenum, New York.

Wegner, M., and Riethmacher, D. (2001). Chronicles of a switch hunt: *gcm* genes in development. *Trends Genet.* **17,** 286–290.

Wu, M. M. J., Cassidy, J. R., and Pukkila, P. J. (1983). Polymorphisms in DNA of *Coprinus cinereus. Current Gen.* **7,** 385–392.

Yamanaka, A., Hayakawa, Y., Noda, H., Nakashima, N., and Watanabe, H. (1996). Characterization of polydnavirus-encoded mRNA in parasitized armyworm larvae. *Insect Biochem. Mol. Biol.* **26,** 529–536.